Supporting Information

Becker *et al.* 10.1073/pnas.0808116105

SI Materials and Methods

Immunoblots. Proteins from severe acute respiratory syndrome (SARS)-coronavirus (CoV)- and Bat-SARS-CoV (S) receptor-binding domain (RBD)-infected cell lysates were resolved and transferred to PVDF as previously described (1). Rabbit polyclonal antisera specific for SARS-CoV nsp1 and nsp8 (2) were precleared on methanol-fixed uninfected Vero cells, protein-A affinity-purified (HiTrap, GE), and diluted to 1 mg/mL for application to blocked membranes at 1:100 dilution for 1 h at room temperature. Rabbit polyclonal antisera specific for SARS-CoV nsp9 and nsp10 (2) were used at 1:500 dilution. Membranes were washed three times for 20 min at room temperature with PBS-0.8% Tween. Goat rabbit-specific antibodies conjugated to horseradish peroxidase (Promega) were diluted 1:2,500 in blocking buffer and incubated with the membrane for 1 h at room temperature and then washed. Labeled proteins were detected by chemiluminescence using Western Lighting Chemiluminescence Reagent Plus (Perkin–Elmer).

1. Graham RL, *et al.* (2005) The nsp2 replicase proteins of murine hepatitis virus and severe acute respiratory syndrome coronavirus are dispensable for viral replication. *J Virol* 79:13399–13411.

2. Prentice E, *et al.* (2004) Identification and characterization of severe acute respiratory syndrome coronavirus replicase proteins. *J Virol* 78:9977–9986.

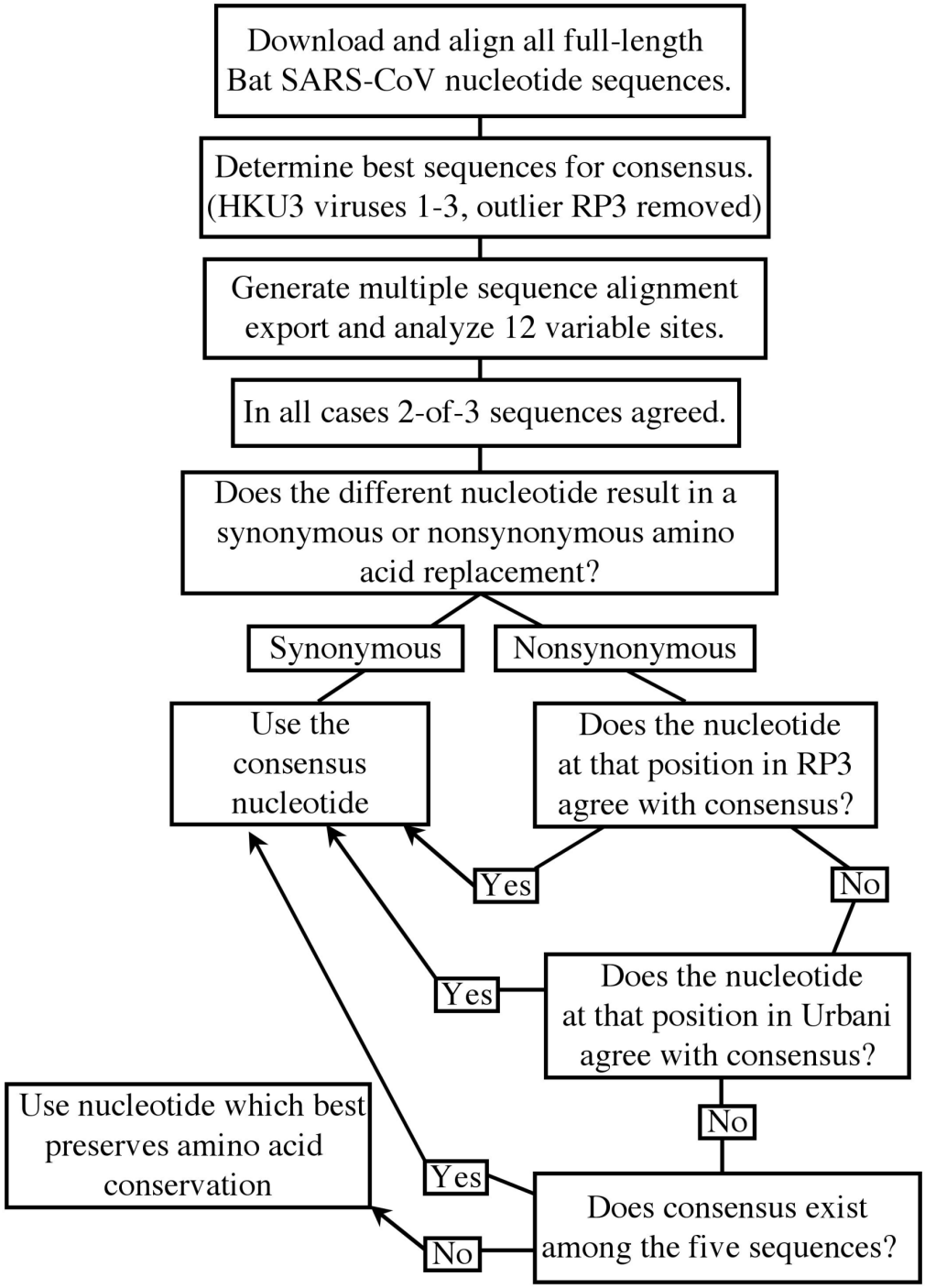

Fig. S1. Determining consensus between the Bat-SCoV sequences. Four full-length Bat-SCoV sequences [HKU3–1 (DQ022305), HKU3–2 (DQ084200), HKU3–3 (DQ084199), and RP3 (DQ071615)] were aligned by using ClustalXv1.83. The 3 HKU sequences shared 99.936% sequence identity (19 changes between all 3), whereas the RP3 sequence had only 89.7% identity with the HKU sequences. SARS-CoV shared 88.3% sequence identity with the HKU viruses, and 92.7% identity with RP3. These analyses indicated that the 3 HKU viruses were the most appropriate for determining a consensus sequence for the 3′ ≈10 kb as only 7 differences were noted between the HKU sequences in the 3′ end of the genome (Fig. S4). Because the 3′ end of CoV genomes encode overlapping ORFs, the nucleotides at these positions in RP3 and SARS-CoV were also included for determining consensus. In all cases, a minimum of 3 of 5 nucleotides agreed at each position and the dominant nucleotide was selected as the consensus residue. For the 5′ two-thirds of the Bat-SCoV genome, 2 of the three HKU sequences agreed at each position. Therefore, consensus was determined by simple majority for nucleotide differences that would result in a synonymous amino acid replacement. If the variation would result in a nonsynonymous change, the nucleotide at the same position in RP3 was included in the analysis, and if this resulted in a tie, the nucleotide at the same position in SARS-CoV was included. The predominant nucleotide at each position was selected as the consensus residue.

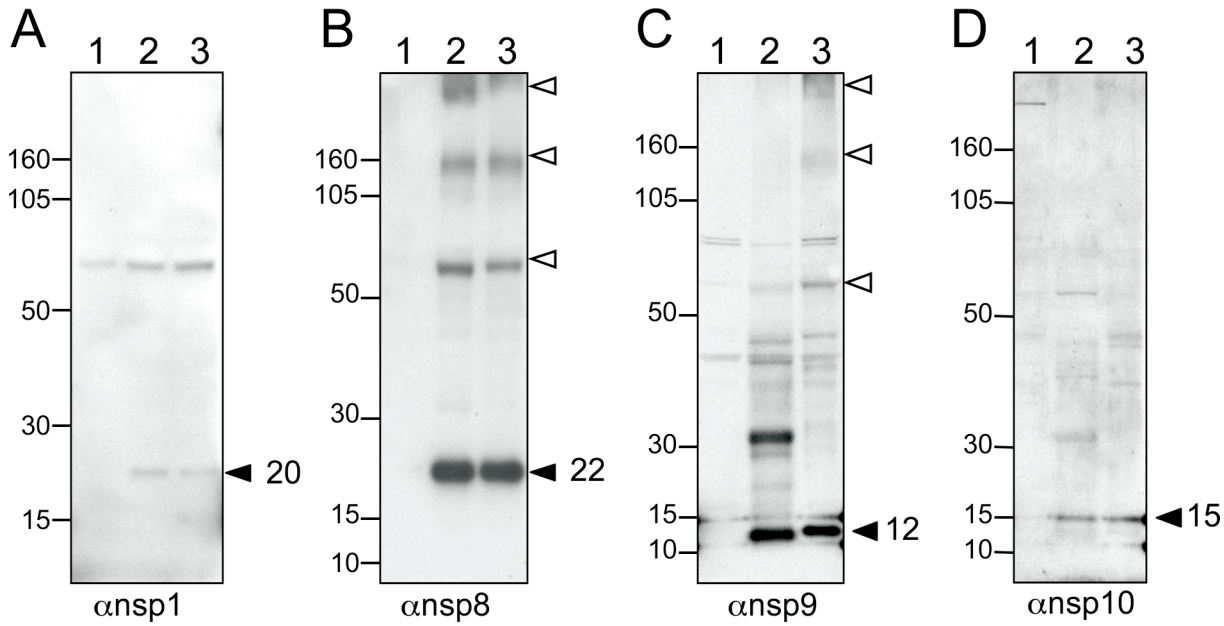

Fig. S2. Comparison of SARS-CoV and Bat-SRBD replicase proteins. Equivalent amounts of mock-infected (lane 1), SARS-CoV-infected (lane 2), or Bat-SRBD-infected (lane 3) cell lysates were probed with rabbit polyclonal Abs specific for SARS-CoV nsps 1, 8, 9, or 10. Relative molecular mass is indicated to the left. Solid arrowheads (◀) to the right indicate the specific proteins detected and the calculated mass, and open arrowheads (◁) indicate possible higher molecular mass polyprotein intermediate precursors.

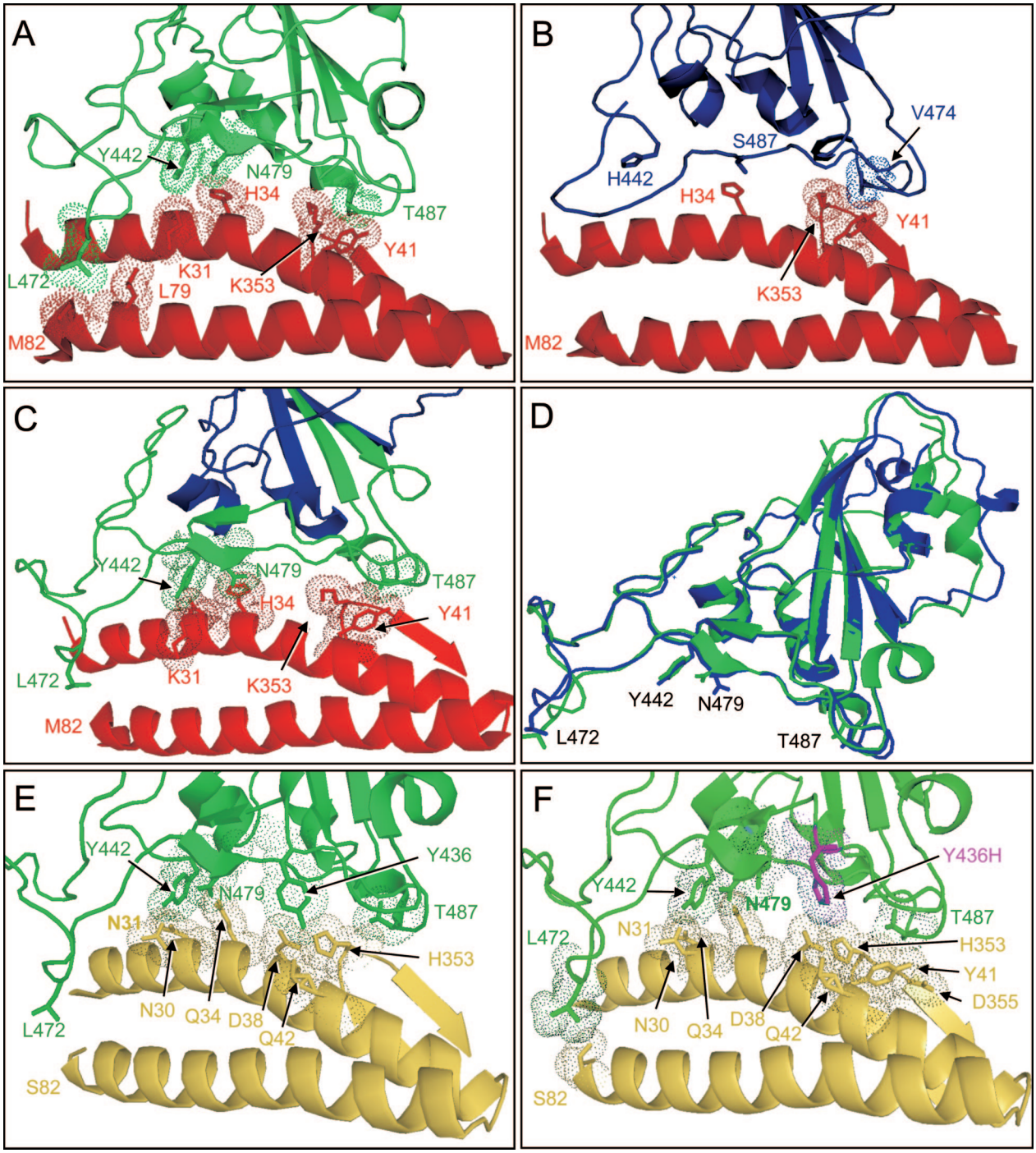

Fig. S3. Modeling interactions of SARS-CoV and Bat-SCoV mutants with human and mouse ACE2. (*A*) The structure of the SARS-CoV RBD (green) interacting with the hACE2 receptor (red) has been solved and shows that 4 residues from the RBD (Y442, L472, N479, and T487) interact with residues of hACE2 to facilitate binding. Mutant viruses that contain full SARS-CoV RBDs, including Bat-F and Bat-SRBD, produce nearly identical models as the one presented here. (*B*) The Bat-SCoV RBD (blue) is significantly different from that of SARS-CoV, with no conserved residues in similar positions, predicting that Bat-SCoV could not interact with hACE2 (red). (*C*) A model of Bat-SRBM (receptor binding motif), in which Bat-RBD (blue) contains SARS-CoV residues in the RBD-hACE2 interface (green). The model predicts that interacting residues are present in similar conformation as in SARS-CoV, but L472 does not interact with M82 and L79 of hACE2. (*D*) Superimposing Bat-SRBM (blue) on the SARS-CoV RBD (green) predicts that regions outside the RBM and distal to the interaction interface are significantly different. (*E*) Modeling the interaction of SARS-CoV RBD (green) with the mACE2 receptor (gold) suggests that differences in mACE2 would reduce or eliminate binding with the SARS-CoV RBD (green). (*F*) Modeling the SARS-CoV MA15 RBD (green) interacting with mACE2 (gold) predicts that the Y436H substitution in MA15 (magenta) will restore binding of Y442, L472, N479, and T487. Also, the interaction of residue T487 (MA15) was increased to include residues D355 and Y41 (mACE2). Residue 436 is predicted to interact with D38 and Q42 of mACE2 in both cases; however, 436H is predicted to interact with additional residues G482 and F483 (not seen in model). Homology models of PDB file 2AJF were constructed by using Modeller and visualized with MacPyMol and Chimera.

	23018	23422	23427	24932	27275	28207	28986
HKU3-1	C	C	C	A	A	A	T
HKU3-3	C	C	C	A	A	A	T
HKU3-2	T	T	T	C	G	G	C
RP3	T	C	T	C	A	G	C
SARS-CoV	T	C	T	C	A	G	C
Consensus	T	C	T	C	A	G	C

Fig. S4. Differences between the 3′ end of the four Bat-SCoV sequences. Seven sites differed between HKU3–1, HKU3–2, and HKU3–3 sequences from the beginning of the Spike gene through the end of the genome. The nucleotides at these positions in SARS-CoV and RP3 were also used to determine consensus at each position. Gray boxes mark the consensus residue for each column. Numbering is based on the HKU3–1 sequence.

Engineered Sites	nt position	HKU3-1	HKU3-3	KHU3-2	Bat-SCoV	RP3	SARS-CoV	Type of Change	Consensus aa	Nonsynon aa
	602	T	T	C	T			syn		
	1837	C	C	T	C	A	A	nonsyn	Thr	Iso
	2827	T	T	A	T	A	A	nonsyn	Val	Glu
Marker mutation	2831	T	T	T	A			syn		
	4280	C	C	T	C			syn		
Adds *Bgl*I site for A/B junction	4385	T	T	T	C			syn		
	6251	A	A	G	A			syn		
	7724	C	C	T	C			syn		
Adds *Bgl*I site for B/C junction	8723	T	T	T	C			syn		
	16043	T	T	C	T			syn		
	16946	A	A	G	G	G		nonsyn	Asp	Asn
	17011	A	A	G	A			syn		
Adds *Bgl*I site for C/D junction	18931	A	A	A	C			syn		
Adds *Bgl*I site for C/D junction	18932	T	T	T	C			syn		
Adds *Bgl*I site for C/D junction	18934	G	G	G	C			syn		
Adds *Bgl*I site for C/D junction	18937	T	T	T	G			syn		
	19729	A	A	C	A			syn		
	19791	C	C	T	C	C		nonsyn	Ala	Val
	20276	C	C	T	C	C		nonsyn	His	Tyr
Knocks out *Afl*II site	21349	G	G	G	A			syn		
Adds *Afl*II site for E1/E2 junction	21374	T	T	T	C		C	nonsyn	Leu	Phe
SARS/Bat-SCov linker	21376	A	A	A	T		T	syn		
SARS/Bat-SCov linker	21379	A	A	A	G		G	syn		
SARS/Bat-SCov linker	21397	C	C	C	T		T	syn		
SARS/Bat-SCov linker	21406	C	C	C	T		T	syn		
SARS/Bat-SCov linker	21412	G	G	G	T		T	syn		
SARS/Bat-SCov linker	21415	C	C	C	G		G	syn		
SARS/Bat-SCov linker	21421	G	G	G	A		A	syn		
SARS/Bat-SCov linker	21448	T	T	T	C		C	syn		
SARS/Bat-SCov linker	21457	A	A	A	G		G	syn		
SARS/Bat-SCov linker	21460	C	C	C	T		T	syn		
SARS/Bat-SCov linker	21478	C	C	C	T		T	syn		
Adds *Age*I and *Bsa*WI sites	22138	G	G	G	C			syn		
Adds *Bst*BI site	22234	T	T	T	G			syn		
Adds *Bst*XI site	23011	A	A	A	G			syn		
	23041	C	C	T	T	T	T	syn		
Adds *Bsi*Wi site	23137	A	A	A	C			syn		
Adds *Hpy*188III site	23156	T	T	T	C			syn		
	23445	C	C	T	C	C	C	nonsyn	Ser	Phe
	23450	C	C	T	T	T	T	syn		
Adds *Bgl*I site for E/F junction	24016	A	A	A	C			syn		
Adds *Bgl*I site for E/F junction	24022	A	A	A	G			syn		
	24955	A	A	C	C	C	C	syn		
	27298	A	A	G	A	A	A	nonsyn	Gln	Arg
	28230	A	A	G	G	G	G	syn		
	29009	T	T	C	C	C	C	nonsyn	Thr	Ile

Fig. S5. Comparison of the differences between the Bat-SCoV and SARS-CoV sequences. This figure outlines the strategy used to design the full-length Bat-SCoV infectious clone. Consensus positions are noted, with only bases used to determine consensus shown. Engineered sites are also shown, with the corresponding alteration noted at the top. Nucleotide position numbering is based on the Bat-SCoV clone sequence, which includes 23 additional nucleotides at the 5′ end and 47 additional bases at the 3′ end as compared with HKU3–1. The SARS-CoV/Bat-SCoV linker region is highlighted in black. The linker is identical at the amino acid level between SARS-CoV and HKU3–1, and 92% identical at the nucleotide level. Synonymous versus nonsynonymous amino acids are noted to show that at some positions there was heterogeneity in the peptide sequence between viral sequences. In these cases, RP3 and, if necessary, SARS-CoV sequences were used to provide further support for consensus. Type of change: syn, synonymous; nonsyn, nonsynonymous.

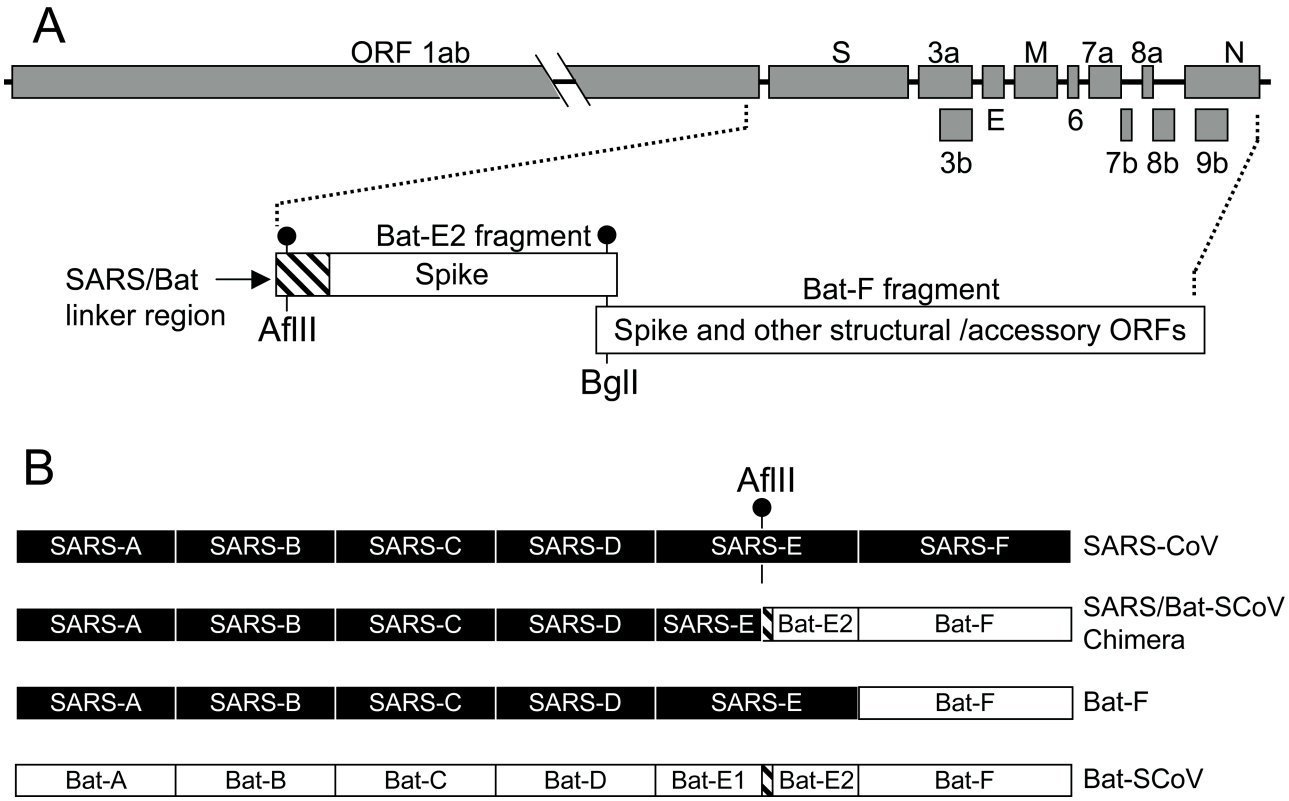

Fig. S6. Designing SARS-CoV/Bat-SCoV chimeric viruses and the full-length Bat-SCoV infectious clone. (*A*) Genome organization of SARS-CoV is shown. The 3′ end of Bat-SCoV, including all of the structural and accessory ORFs, was divided into 2 fragments called Bat-E2 and Bat-F, and a unique BglI restriction site was engineered into each fragment by using the SARS-CoV fragment as template. A 114 nt linker region identical to the SARS-CoV nucleotide sequence was added at the 5′ end of the Bat-E2 fragment (hatched box), to incorporate a unique AflII restriction site. This was designed so that the SARS-E and Bat-E2 fragments could be digested with AflII and ligated so that gene 1 would originate from SARS-CoV, and genes 2–9 would originate from Bat-SCoV. However, the linker region is strictly conserved at the amino acid level, indicating that nsp16 function would not be altered. (*B*) The Bat-SCoV infectious clone was designed using the SARS-CoV clone as a template. BglI restriction sites were engineered into the sequence to match those of SARS-CoV, allowing for exchange of fragments. The constructs include: SARS-CoV, the Urbani infectious clone; SARS/Bat-SCoV chimera, incorporating SARS-CoV gene 1 (ORF1ab) and Bat-SCoV structural and accessory ORFs (genes 2–9); Bat-F, comprises of SARS-CoV gene 1 and part of gene 2 including the RBD, and Bat-SCoV from the 3′ end of gene 2 through gene 9; and the Bat-SCoV clone.

Table S1. Nucleotide variation in Bat-SRBD p2 population virus

Wt nt	Protein	Wt aa	Alternative nt	Alternative aa
5962 A	nsp3	1082K	G	E
10561 A	nsp5	199T	C	P
25568 T	ORF3	119S	G	R
27720 A	ORF7 5′ UTR	—	G	—

UTR, untranslated region.

Table S2. Comparison of SARS-CoV and Bat-SCoV predicted proteins

Mature protein	Residues	Identical, %	Changes	Similar
nsp1	180	98	3	3
nsp2	638	90	61	39
nsp3	S1922/B1916	91	158	78
nsp4	500	95	23	14
nsp5	306	99	2	2
nsp6	290	98	3	3
nsp7	83	98	1	1
nsp8	198	98	2	2
nsp9	113	99	1	1
nsp10	139	99	1	1
nsp12	932	98	12	10
nsp13	601	99	5	4
nsp14	527	98	10	7
nsp15	346	97	7	4
nsp16	299	97	6	3
ORF2 (S)	1262	78	275	108
ORF3a*	275	81	50	24
ORF4 (E)	77	100	0	NA
ORF5 (M)	222	98	3	3
ORF6	64	90	4	4
ORF7a	123	94	7	4
ORF7b	45	93	3	3
S:ORF8a/B:ORF8†	S39/B121	12	24 of 39	9
S:ORF8b/B:ORF8†	S84/B121	18	62 of 84	12
ORF9 (N)	423	96	13	6

S, SARS-CoV; B, Bat-SCoV; NA, not applicable.

*SARS-CoV encodes 2 overlapping ORFs within this subgenomic mRNA, called ORF3a and ORF3b. Bat-SCoV has an early stop codon and encodes for a single protein in ORF3 which corresponds to the SARS-CoV ORF3a.

†Bat-SCoV encodes for 1 protein in ORF8, while SARS-CoV encodes 2 overlapping proteins called ORF8a and ORF8b.

Table S3. Primers used for generation of Bat-SCoV/SARS-CoV receptor binding residue substitutions

Primer name	Sequence	Used to amplify	Sense
Bat A	5′-CCTTTGGAATTAATATCACTTCTTATAGAGTAGTTATGG	Bat-E2	+
Bat RBD B	5′-GCATAGCGTCTCCATGTTAGGAAATCTAATAACCTCTTG	Bat-E2	−
Bat RBM B	5′-GCATAGCGTCTCCCTAGTATTCCAAGCAATTACACAGC	Bat-E2	−
Urb RBD C	5′-GCATAGCGTCTCTACATTACAAACTTGTGTCCTTTTGGAG	SARS-E	+
Urb RBM C	5′-GCATAGCGTCTCCCTAGGAACATTGATGCTACTTCAACTG	SARS-E	−
Urb D	5′-GCATAGCGTCTCCAGGTCAGTGGATAATTTTGGTCCACAAA	SARS-E	−
Bat E	5′-GCATAGCGTCTCGACCTGGTTAAGAACCAGTGTGTTAATTTC	Bat-E2	+
Bat F	5′-CATAAGAGGCGTTGACATGCTCAGCTCCTATAAGAC	Bat-E2	−
Bat Motif G	5′-GCATAGCGTCTCCCCCTTGACTACGAATGTATCAGCATACACACTTG	Bat-E2 RBM	−
Bat Motif H	5′-GCATAGCGTCTCCAGGGAGATGATGTAAGACAAGTTGCACCAGGTGAAAC	Bat-E2 RBM	+

Underlining indicates residues incorporated for cloning or mutagenesis purposes.